APPENDIX A

KIEV INTERNATIONAL INSTITUTE OF SOCIOLOGY

Ukraine SME Survey

METHODOLOGICAL REPORT

Kiev, 1999

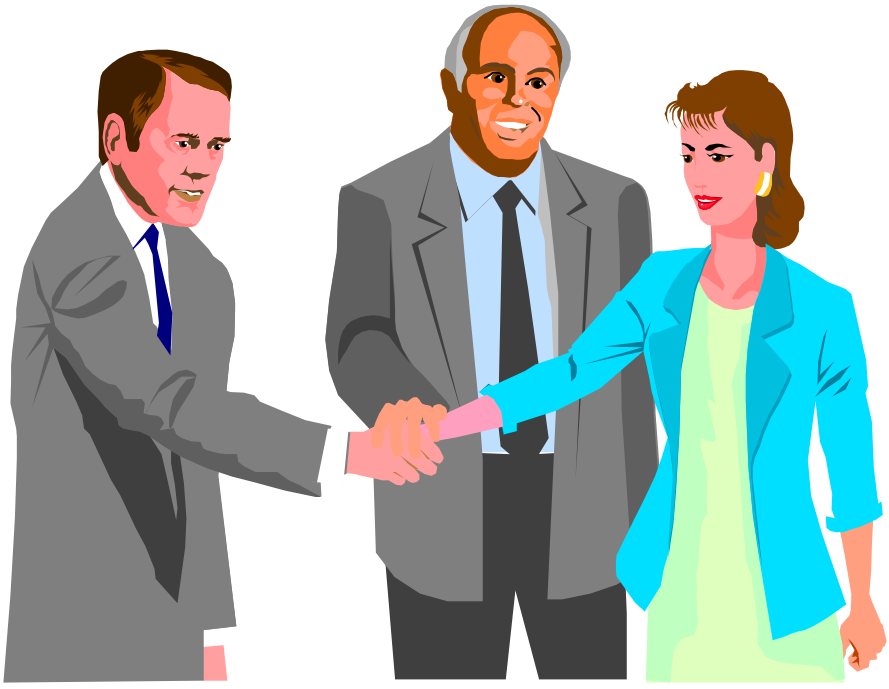

TABLE OF CONTENTS

, in which we fixed the social-demographic information about all household members aged 15 and older and on forms of their employment (wage employees or self-employed). For polling hired employees we used the abbreviated version of questionnaire for

employers which included only such questions about the enterprise-employer in which the wage employee was quite competent.

On making street survey of visible firms we used the **"Rout Map of Businesses"** to which we inserted additional information about the firms selected at random (name, kind of activity, number of employed).

In total, the final data file of the survey included 3904 interviews of registered enterprises, 3267 interviews of wage employees, 646 interviews of self-employed household members, and also 996 interviews of "visible" firms during street polling.

The response rate in three directions of survey made:

- questioning registered firms– 67.5%
- questioning households- 81.2% for households and 87.5% for household members.
- street questioning of businesses (enterprises) – 66.9%

The sampling error in questioning households does not exceed 2.6%.

1. DEVELOPMENT OF TOOLS

1.1 Preparation of questionnaires.

Development of tools included preparation and pre-test of questionnaires, Manuel for the interviewer on making polling, and other field documents (see full set of field documents in SUPPLEMENT).

1.2 Pre-test.

The purpose of pre-test is the description of possible problem situations by means of modelling of survey process "in miniature" and the development of proposals for elimination of possible problems at all stages of further work.

In this pre-test the following tasks were set before the project task force:

- Testing the procedure of sampling filing;
- Testing accessibility of potential respondents;
- Testing field tools of survey (questionnaire, directions to interviewee on respondents selection and work with questionnaire);
- Evaluation of terms of field work fulfil (time spent by each interviewer for making a certain number of interviews)

- Testing the procedure of field control and data input
- Evaluation of compliance of data received during the polling to tasks of survey.

The pre-test was held in February 9-17 and included:

- preparation and edition of questionnaire in Ukrainian and Russian languages
- formation of enterprise sampling
- Briefing for interviewers on selecting respondents and work with questionnaire
- Pre-test data input into computer
- Holding briefing on the pre-test results with the participation of KIIS research group and interviewers.
- Preparation of report on pre-test on the basis of interviewers reports and analysis of marginal.

Five experienced interviewers from KIIS who had a special training for carrying out pre-test and participated in practically all KIIS pre-tests for recent 3-5 years took part in this pre-test.

The field stage of the pre-test was held in three directions:

1) Polling of enterprises on the basis of address sampling. In this case the interviewers received the addresses of enterprises in advance for polling (questioning).

2) Street polling of “visible” enterprises, selected for polling at random on the territory of postal districts determined in advance.

3) Polling in the households of all employed household members aged 15 and older.

In total, 57 interviews were held during pre-test, of them 26 - at enterprises from address sampling, 9 - at enterprises from street sampling. During the polling in 16 households 20 people were interviewed, 18 of them had 1 job, and 2 persons - 2 jobs. That's why the number of questionnaires collected from these households made 22.

The main conclusions from pre-test are, first, approximate definition of response rate for polling registered enterprises according to address lists. On estimation of interviewers and KIIS research group, the expected response rate had to be about 30-40%. The second conclusion from pre-test was the understanding of the fact that before the beginning of the field stage a detailed terminology training for interviewers of this survey was needed - in particular, they should learn the main economical categories used in this questionnaire.

Besides, the pre-test showed the problems of questionnaire and became the basis for the revision of tools and manuals for questionnaire.

2. SAMPLING

2.1 Sampling development

As it was mentioned above, the survey included three directions based on different principles of sample:

1) Polling of 5000 registered Ukrainian enterprises;

2) Polling of 4000 households;

3) Street polling of “visible” businesses.

2.1.1 Sampling of registered enterprises.

Sampling of registered enterprises was formed on the basis of lists of State Committee for Statistics of Ukraine, which has several databases of ALL registered enterprises. This list was formed on the basis of State Committee for Statistics own files (each newly created enterprise undergo an obligatory registration at the State Committee for Statistics), Tax Inspection files (Tax Inspection passes the current information on closed enterprises to the State Committee for Statistics), and also municipal bodies files.

The most complete database of State Committee for Statistics accounts for about 600000 enterprises. Entry on each enterprise includes its name, registration code, address, type of activity, form of ownership. Besides, there exists the database which contains 250000 enterprises with the indication of the employment size (number of employed) at each enterprise.

The main principle of sample of the registered enterprises was the quoting of enterprises by the employment size, that’s why the main body of sampling (enterprises with the employment size more than 10 people) was formed on the basis of the second data base (250000).

At the preliminary stage of survey (February 1999) KIIS made an agreement with the Marketing Department of the Research Institute of Statistics of the State Committee for Statistics of Ukraine on acquiring sampling on the basis of databases 250000 and 600000 (see the KIIS complete request for sampling formation in SUPPLEMENT).

The sampling was prepared according to the following characteristics:

À. On the basis of data base 250000:

1) 3000 enterprises with number of employed 251 and more;

2) 3000 enterprises with number of employed from 51 to 250;

3) 3000 enterprises with number of employed from 11 to 50;

B. On the basis of database 600000 - 6000 enterprises, among which only the enterprises with the number of employed up to 10 people were to be polled.

The last sample made up of 6000 enterprises was based on the basis 600000 because of the following reasons:

à) as the database 250000 consisted of enterprises regularly reporting to the corresponding state bodies, it was possible to select from database 600000 the sub-group of enterprises that were not very stable and, thus, were of special interest for our survey (we regarded the hypothesis that the share of informal activities is quite large at such enterprises);

b) as the small enterprises are considered to be the most unstable group of enterprises, the decision was taken that on the basis of base 600000 the sampling of enterprises with the least number of employed (up to 10 people) would be formed.

On forming all four sub-samplings the following criteria were considered:

- The enterprises which are not engaged in production and sale of goods in Ukrainian market were excluded (budget organisations - state schools, hospitals, kindergartens; public and religious organisations, state bodies and power structures, etc.); also were excluded from the list the enterprises engaged in agricultural activity with the exception of enterprises, which besides the agricultural activity are engaged in transportation and processing of agricultural production.

- The enterprises which are engaged in agricultural activity were driven out of the list with the exception of enterprises which besides agricultural activity are engaged in transportation and processing of agricultural production.

- All enterprises were selected in 200 populated areas, which corresponded to the populated areas where the household polling was carried out.

Each enterprise from three groups of employment size on the basis of base 250000 received a random number with the help of random number generator according to which the enterprises were sorted out. The first 3000 enterprises from each group made a sampling of enterprises necessary for survey.

All enterprises of the base 250000 where the number of employed exceeds 10 persons were excluded from sampling on basis 600000. As a result, we obtained the database of enterprises necessary for the 4^{th} sampling of 6000 enterprises. Each enterprise of this base with the help of random number generator also received a random number according to which this base was sorted out. The first 6000 enterprises were included into sampling.

On estimation of the State Committee for Statistics experts, the maximally possible mistake in final files containing the information on 15000 enterprises makes 10%. This mistake is the result of incorrect or wrong entries in the primary bases of State Committee for Statistics

of Ukraine (according to the results of questioning which was held on the basis of this sampling, this error appeared to be larger - see Chapter 4 of this report).

Thus, KIIS received a sampling consisting of 15000 enterprises of which 5000 enterprises were to be polled. In other words, this sampling contained a triple supply of enterprises for polling.

Further, before the beginning of fieldwork the sampling on four groups of enterprises was divided into separate lists according to the number of brigades of questioning network of interviewers with preserving of ordering at random numbers. Further, the lists were divided into 3 equal parts - the main list, the first supplementary, and the second supplementary. According to the conditions of realising this sampling by the interviewers they were forbidden to pass to another list before all the addresses from the previous list would have been worked out.

2.1.2. Sampling development on households polling

The sampling which was used in household's survey is the standard KIIS procedure worked out on the basis of method of constructing survey for polls representative for the population of Ukraine.

2.1.2.1. Sampling description

In conducting this research over the past nine years, we have been able to develop a survey centre that can claim to approach international standards in survey research. We have participated in an eighteen-month international program on sample design, with the participation of the world-renowned sampling expert Leslie Kish and the chief of sampling for the Survey Research Centre of the University of Michigan, Steve Heeringa; both have contributed specific suggestions to our sample design, which is still evolving (see Swafford M., Kosolapov M., Kish L., Heeringa S.. Sample Design for Republics of the Former Soviet Union. - National Council for Soviet and East European Research, 1995; that book has also the description of our sample, but not the last one).

A specific feature of our approach to sampling is ensuring random selection at every stage of sampling.

2.1.2.2. General concept of the sampling

We, as usual, tried to use a design close to a true probability sampling, since only that "guarantees" unbiased results and the proper calculation of sampling errors. This means that we shun quota samples[1], snowball techniques, random-walk samples, haphazard street interviews,

[1]We realize that some reputable Western marketing organizations use quota samples to cut costs. However, presumably they have taken the trouble to conduct methodological studies to demonstrate that their approach produces the same results as the standard produced by a probability sample. No such methodological studies have been published about quota (or random-walk) samples in the NIS, and it would be naive to assume that the results would be the same. This is especially true at the last stage of selection.

and any other methods that permit either discretion on the part of interviewers or self-selection by respondents.

When it goes about sampling within a populated entity in the NIS, most survey organisations use address bureaus or voter lists as their sample frames. The disadvantage of lists based on the address bureau is that, especially in large cities, a sizeable proportion of people do not live where they are registered. Also, certain categories of people -- so-called limitchiki, whose registration depends on holding (often undesirable) employment in the city -- may be represented improperly. Furthermore, lists are simply not as well-maintained as they were under the Soviet regime. Housing lists could in principle be used as a basis for building a list of dwellings and apartment numbers (rather than household identities) to serve as a sample frame. However, the information is contained on hand-written cards, which are more awkward to use than alternatives.

As for voting lists, they do not include individuals under eighteen, so they do not serve well in a study of those between 15 and 18. Again, they are seemingly less well-maintained than they were under the Soviet regime, when activists were charged with getting more than 99.5% of the population to the polls. Without special provisions, households with more voters would be more likely to fall in the sample than those with one voter. Finally, the lists are maintained in hand-written form in many locations. So again, it would be awkward, though possible, to use the lists merely as a basis for building a sample frame of addresses or households.

Because of these drawbacks, KIIS has for several years pioneered the effort to adapt Western area-sample techniques to the situation in the Ukraine. In doing so, it has of course had to work around the lack of certain statistics that are routinely available in the United States, such as block or census tract statistics.

Here is the general outline of our approach. We use settlements as primary sample unit's (PSU's) and use probability, proportional to the size to select it. Then we select randomly postal districts in every selected settlement. Addresses or individual dwellings constitute the next sampling units. The boundaries of postal districts are well defined, and nobody knows the nooks and crannies in a postal district better than letter carriers. Hence, we escape the objection of using questionable lists and the need to seek inaccessible data.

On the territory of each postal district we randomly select streets, houses and apartments. After that we conduct a survey itself - selection and questioning of the respondents (we use special procedure - see below - to avoid bias in respondent selection). If after selection of the respondent it turns out that the respondent is not at home, the interviewer will make up 2 call-backs in order to find the respondent at home.

2.1.2.3. Sample size

Since the target sample size is 4000, following Western practice, we would draw a larger sample (about 4200) in anticipation of modest non-response. We do not employ the common practice in the NIS of substituting households with characteristics similar to those who fail to participate for some reason. The most important procedure is to keep selection random to avoid bias. This requires repeated call-backs to interview difficult-to-reach households whose work

lives may well differ systematically from those who are easily reached. Even if one knows the attributes of difficult-to-reach households, substituting an easy-to-reach household on the basis of size and structure accomplishes nothing salutary if that new household tends to differ from the difficult-to-reach household by virtue of the fact that the new household is easier to reach. Besides interviewer himself is interested in increasing of part of the households which took part in the interviews, as he is paid only for the interviews actually conducted. Information about the portion of households been interviewed makes possible to estimate the sample mistake more exactly (in case of substitution this mistake is camouflaged by the easy-to-get households and causes an uncontrolled shift in answers).

2.1.2.4. First stage of the sample – settlement selection.

Ukraine is divided into 24 oblasts and the Crimea. One city in each of these 25 administrative entities is designated as the o*blast centre*. Further, each of the 25 administrative entities is divided into *raions*, rather like counties in the United States and Britain.[2]

In principle, a national sample need not contain households from every oblast. However, our sample is in effect stratified by oblast. More precisely, the sample is drawn in such a fashion that every oblast is *proportionally* represented. This increases the cost of fieldwork because it requires that we maintain a staff in all oblasts, but it greatly enhances our ability to characterise regions, not just the country as a whole.

All households in each oblast are officially considered to live in one of three kinds of populated areas: cities, "villages of the city type" (PGTs), and villages. The designation correlates largely with size; however, it ultimately refers to the type of municipal administration granted to the settlement. Thus, it turns out that a few PGTs are larger than some small cities; and some villages are larger than some PGTs.

In official statistics, the rural population consists only of villages; the urban population is defined as comprising all cities and PGTs. The sample is stratified according to whether population points are considered urban or rural. The urban and rural populations of each oblast are represented proportionally.

We select PSU, using PPS (probability, proportional to the size).

We can draw samples very quickly because we have computerised the procedure. We developed our own software for sample design. For the same reason, it is quite easy for us to change the parameters of the sample to meet special needs.

---

[2]To avoid repeatedly writing "the 24 oblasts and the Crimea", we shall use the term "oblasts" henceforth to refer to all 25 administrative entities. Strictly speaking, however, the Crimea in not an oblast, and referring to it as such would be considered a political *faux pas* by many.

2.1.2.5. Stratification of households on administrative-territorial units of Ukraine and inside them on urban and rural population.

Statistic grounding of the way of stratification on the number of population is based on the following positions:

Existing statistics in Ukraine does not allow to directly evaluate the quotes of households for the necessary categories of population. Existing statistics allows to represent the general aggregate in terms of different categories of families. Family statistics allows to assert that the number of families on each territorial-administrative category of population with the precision up to statistic significance is proportional to the total number of population within this category. It means that for the evaluation of strata it is possible to use the total number of population of stratified category, and not the number of households which the given category includes.

It should be noted that this polling confirmed the correctness of chosen approach for quoting households.

2.1.2.6. Second stage of the sample – selection of postal districts

Within each PSU, a list of residential postal districts is constructed. Sometimes this can be done centrally; at other times, a fieldworker must consult local authorities to construct the necessary list in a given PSU. Once the list is constructed a few postal districts are chosen randomly taking into account the number of residential routes, which is proportional to the number of apartments. The number of postal units selected depends on the size of the ultimate clusters that we want.

2.1.2.7. Third stage– households selection

Once the postal districts are selected, a rule for designating the first household is developed which randomly selects a carrier route (if there is more than one), a street on the route, a building on the street, and an apartment in that building. From this first apartment, the interviewer should interview each following household, but not more than 10 households.

2.1.3. Sampling development for street survey of "visible" businesses.

Populated areas and post-offices (postal districts) for this type of survey were selected only for urban population of Ukraine similar to procedure of selection of populated areas and postal districts for households' survey. Furthermore, on the territory of each postal district the first 30 "visible" firms situated on the interviewer route were fixed (see rules of interviewer route unwrapping in more details in SUPPLEMENT). The list of firms obtained in such a way due to random procedure of selection is the representative sampling of "visible" enterprises situated in Ukrainian cities.

The next step of selection was the purposeful selection of firms with the least number of employed from lists of enterprises, which had already been compiled. The small firms made the greatest interest in this part of research, as we would not be able to obtain a reliable information on the activity of non-registered enterprises from any of two other directions of research. In research of firms on the basis of State Committee for Statistics sampling only registered

enterprises were questioned; in research of households the share of people who sincerely gave answers about such activity didn't reflect the real degree of non-registered enterprises prevalence. In the countryside the questioning on random sampling was not held as in the majority of Ukrainian settlements the number of enterprises is not sufficient for filling each PSU (Primary Sampling Units).

2.2. Evaluation of sampling quality.

In total, the datafile contains the results of 5546 interviews with subjects of business (entrepreneurship) - 3904 interviews with the managers of enterprises registered in Ukraine selected on the basis of State Committee for Statistics data, 646 interviews with self-employed household members, and 996 interviews with the managers of enterprises selected by the route way in the streets of Ukrainian cities and towns. Chapter 4 of this report gives details of problems of sampling realisation. In the current chapter we'd like to dwell upon the evaluation of quality of realised sampling, basing mainly on the comparison of distributions of different characteristics of aggregate received in the framework of this survey research with existing statistic data on corresponding characteristics of general population.

2.2.1. Address sampling of registered enterprises.

The theoretical evaluation of quality of sampling presented to KIIS by the State Committee of Ukraine for Statistics can be obtained by means of comparison of statistic data of enterprises distribution according to spheres of their activity (on database - 250000) with similar distribution of enterprises according to the data obtained during research. The most unambiguous characteristics according to which it is possible to evaluate the differences between general and sample population are the sphere of activity of enterprises. That's why only this characteristic will be used in this chapter (the form of enterprise ownership could not serve as such characteristic as the State Committee for Statistics used more than 50 codes of forms of ownership of enterprises which are impossible to be correlated with similar information obtained during the questioning).

Table 2.2.1 Comparison of distribution of data from Sate Committee for Statistics data base - 250 000 with data of research "Ukraine MSE Survey" according to spheres of activity of enterprises inside each group of employment size

Spheres of activity	Employment size 1-10, %			Employment size 11-50, %			Employment size 51-250, %			Employment size 251 and more, %		
	Database	Data of survey	**Difference**	Database	Data of survey	**Difference**	Database	Data of survey	**Difference**	Database	Data of survey	**Difference**
Construction	9.7	6.2	**-3.5**	15	12	**-3**	22.1	25.2	**3.1**	11.3	11.1	**-0.2**
Industry	13.5	9.3	**-4.2**	18	13.8	**-4.2**	30.7	22.6	**-8.1**	57.3	50.5	**-6.8**
Trade and public catering	52.8	42.4	**-10.4**	40.3	38.3	**-2**	19.6	18.2	**-1.4**	6.8	7.4	**0.6**
Transport and communication	1.4	3.3	**1.9**	2.2	3	**0.8**	6.3	6.7	**0.4**	10.7	12.6	**1.9**
Finances, credit, insurance, provision of pensions	1	2.5	**1.5**	1.9	1.3	**-0.6**	1.5	2.1	**0.6**	2.1	2	**-0.1**

Agriculture and forestry, storage	0.9	2.5	**1.6**	1.3	2.4	**1.1**	2.5	3.9	**1.4**	0.3	3.9	**3.6**
Housing-municipal management, material and technical supply, hotels, recreation	6.7	10.1	**3.4**	10.2	13.4	**3.2**	9.8	8.1	**-1.7**	5.7	6.4	**0.7**
Social and cultural services (education, health protection, etc.)	3.6	5.6	**2**	4.4	7.2	**2.8**	3	4.4	**1.4**	1.7	1	**-0.7**
Science and services of science	5.2	6.5	**1.3**	3.3	3.4	**0.1**	3.1	3.2	**0.1**	3.8	2.1	**-1.7**
Other business-services	3.3	4.2	**0.9**	1.2	0.8	**-0.4**	0.3	0.4	**0.1**	0	0	**0**
Other	1.8	7.0	**5.2**	2.1	4.1	**2**	1	4.9	**3.9**	0.3	2.9	**2.6**
Hard to say	-	0.4	**0.4**	-	0.3	**0.3**	-	0.2	**0.2**	-	0.1	**0.1**

As seen from table 2.2.1. the maximal discrepancy of research data with statistic data makes 10.4%. According to us, such discrepancy is dealt with shortage of interviews at the enterprises of the 1st group of employment (read about the reason of this shortage lower). It is worth noting that the trade enterprises with the number of employed less than 10 people are the least stable for today in Ukrainian market. So, to reduce the quantity of mistake the sampling for research of enterprises of such type has to be based on the information of the current year (sampling for this research corresponds to the situation of 1997). That's why the part of enterprises from the address sample does not exist for today, and the new enterprises which were created for this period had not been included into the Ministry for Statistics data base. This fact provoked such a level of a sampling error.

The shortage of large industrial enterprises (sampling error is 8.1% for the 3rd group of employment and 6.8% for the 4th group of employment is dealt with primary displacement of sampling for the enterprises of this group. As seen from table 2, sampling on 200 populated areas used in this research, covers in the 1st and the 2nd groups of employment more than 60% of enterprises of general population, when in the 3rd and 4th groups not more than 40% of all enterprises of corresponding size are concentrated in selected populated areas.

Òàble 2.2.2. Distribution of enterprises of each group of employment size on populated areas under research and other populated areas of Ukraine (statistic information on database 250000):

	Enterprises situated….	
	In 200 settlements under research, %	In other settlements that didn't get into sample, %
1-10 employees	71.70	28.30
11-50 employees	62.24	37.76
51-250 employees	38.92	61.08
251+ employees	33.25	66.75
TOTAL	62.44	37.56

Table 2.2.3 presents primary displacements of statistics of general population of enterprises from database 250000 compared with enterprises situated in 200 selected populated areas.

Òàble 2.2.3. Comparison of data distribution of database of State Committee for Statistics - 250000 on all settlements and on 200 settlements selected for research:

Spheres of activity	Employment size 1-10, %			Employment size 11-50, %			Employment size 51-250, %			Employment size 251 and more, %		
	All settlements	200 settlements	**Difference**	All settlements	200 settlements	**Difference**	All settlements	200 settlements	**Difference**	All settlements	200 settlements	**Difference**
Construction	9.7	9.6	**-0.1**	15	14.5	**-0.5**	22.1	26.1	**4**	11.3	13.6	**2.3**
Production types of everyday services to population	4.2	3.4	**-0.8**	4.8	4.2	**-0.6**	2.6	3	**0.4**	0.8	1	**0.2**
Geology and prospecting, geodesic and hydrometeorology services	0.2	0.1	**-0.1**	0.2	0.2	**0**	0.3	0.4	**0.1**	0.4	0.2	**-0.2**
Housing-municipal management	1.8	1.7	**-0.1**	3.6	3.5	**-0.1**	5	5.4	**0.4**	4	4.8	**0.8**
Joint commercial activity on ensuring functioning of the market	3.3	3.8	**0.5**	1.2	1.6	**0.4**	0.3	0.5	**0.2**	0	0	**0**
Storage	0.9	0.6	**-0.3**	1.3	0.7	**-0.6**	2.5	0.8	**-1.7**	0.3	0.2	**-0.1**
Information-computational services	1.2	1.5	**0.3**	0.6	0.9	**0.3**	0.3	0.5	**0.2**	0.1	0.2	**0.1**
Other types of activity of material and production sphere	1.3	1.5	**0.2**	1.9	2.1	**0.2**	1	1.6	**0.6**	0.3	0.3	**0**
Culture and art	0.9	0.9	**0**	1.5	1.4	**-0.1**	0.7	1.2	**0.5**	0.3	0.3	**0**
Ì àterial and technical supply and sale	0.7	0.7	**0**	1.8	1.4	**-0.4**	2.2	2.3	**0.1**	0.9	1.2	**0.3**
Education	0.5	0.5	**0**	0.9	1	**0.1**	0.8	1.3	**0.5**	0.3	0.3	**0**
Science and services of science	3.8	4.7	**0.9**	2.5	3.4	**0.9**	2.5	3.9	**1.4**	3.3	3.7	**0.4**
Î perations with real estate	0.5	0.5	**0**	0.2	0.2	**0**	0	0.1	**0.1**	0	0	**0**
Health protection, physical culture, and social welfare	2.2	2.3	**0.1**	2	1.8	**-0.2**	1.5	1.3	**-0.2**	1.1	0.7	**-0.4**
Industry	13.5	12.7	**-0.8**	18	15	**-3**	30.7	24	**-6.7**	57.3	51.9	**-5.4**
Òrade and public catering	52.8	52.8	**0**	40.3	40.7	**0.4**	19.6	19.6	**0**	6.8	6.7	**-0.1**
Òransport and communication	1.4	1.5	**0.1**	2.2	1.9	**-0.3**	6.3	5.6	**-0.7**	10.7	11.6	**0.9**
Finances, credit, insurance, provision of pensions	1	1.2	**0.2**	1.9	1.9	**0**	1.5	2.4	**0.9**	2.1	3.2	**1.1**

At the end we want to show the distribution of enterprises of general population (statistics from database - 250 000) according to their size. Sampling of this research was quoted according to this characteristics, but during the extrapolation the results of survey of registered enterprises on Ukrainian market in general we recommend to use the procedure of weighting data file with consideration of proportion of enterprises of each group of employment size.

Òàble 2.2.4. Distribution of enterprises of database - 250000 and enterprises interviewed during research according to their size.

	Statistic data		UKRAINE SME SURVEY, Register's sample		Difference, %
	Frequency	%	Frequency	%	
1-10 employees	98445	55.8	1178	30.2	-25.6
11-50 employees	42188	23.9	875	22.4	-1.5
51-250 employees	25619	14.5	912	23.4	8.9
251 employees and more	10269	5.8	939	24.1	18.3
TOTAL	176521	100.0	3904	100.0	

***Statistics on database-250000. All enterprises not to be interviewed (which do not operate in the market) are excluded from database.

2.2.2. Households sample.

For testing the quality of this sampling realisation we make the check of the data obtained in the survey research with the data of the State Committee for Statistics for 1998 (tables 2.2.6.-2.2.9.) and the data of all-Ukrainian population census of 1989 (table 2.2.5).

Òàble 2.2.5. Comparative characteristics of statistic data and households file: distribution of households according to size.

Size of the household	Statistic data 12.01.89, %	UKRAINE SME SURVEY, household's sample		Difference, %
		Frequency	%	
1 person	18.2	700	17,5	-0.7
2 persons	28.7	1083	27,1	**-1.6**
3 persons	22.1	842	21,0	-1.1
4 persons	19.7	810	20,2	0.5
5 persons	7.0	340	8,5	1.5
6 persons	2.8	159	4,0	1.2
7 persons	1.4	69	1.7	0.3
TOTAL	99.9	4002	100	

Òàble 2.2.6. Comparative characteristics of statistic data and individual data of survey research: distribution of Ukrainian population aged 15 and older according to sex:

Gender	Statistic data, 1998, %	UKRAINE SME SURVEY, Individual data for adults in the household		Difference, %
		Frequency	%	
Male	45.4	4348	44.4	-1
Female	54.6	5441	55.6	1
TOTAL	100.0	9789	100.0	

Òàble 2.2.7. Comparative characteristics of statistic and individual data of survey research: distribution of Ukrainian population aged 15 and older according to age:

Age	Statistic data, 1998, %	UKRAINE SME SURVEY, Individual data for adults in the household		Difference, %
		Frequency	%	
15-19 years old	9.0	749	7.7	-1.3
20-24 years old	8.9	796	8.1	-0.8
25-29 years old	8.4	733	7.5	-0.9
30-34 years old	8.4	784	8	-0.4
35-39 years old	9.7	863	8.8	-0.9
40-44 years old	9.1	894	9.1	0
45-49 years old	8.7	866	8.8	0.1
50-54 years old	5.5	688	7	1.5
55-59 years old	8.1	607	6.2	-1.9
60-64 years old	6.8	857	8.8	2
65-69 years old	6.3	613	6.3	0
70 years old and older	11.1	1339	13.7	**2.6**
TOTAL	100.0	9789	100.0	

Òàble 2.2.8. Comparative characteristics of statistic and individual data of survey research: distribution of Ukrainian population aged 15 and older according to the type of settlement

Type of the settlement	Statistic data, 1998, %	UKRAINE SME SURVEY, Individual data for adults in the household		Difference, %
		Frequency	%	
City, town	68.1	9744	68.9	0.8
Village	31.9	3048	31.1	-0.8
TOTAL	100.0	9792	100	

Òàble 2.2.9. Comparative characteristics of statistic and individual data of survey research: distribution of Ukrainian population aged 15 and older by regions (oblasts).

Region	Statistic data, 1998, %	UKRAINE SME SURVEY, Individual data for adults in the household		Difference, %
		Frequency	%	
Crimea	5.1	473	4.8	-0.3
Kiev (city)	5.3	492	5.0	-0.3
Êiev oblast	3.7	385	3.9	0.2
Vynnitsa	3.7	352	3.6	-0.1
Volyn	2.0	231	2.4	0.4
Dnepropetrovsk	7.5	685	7.0	-0.5
Donetsk	10.3	951	9.7	-0.6
Zhitomir	2.9	255	2.6	-0.3
Transcarpathian	2.4	331	3.4	1.0
Zaporozhie	4.1	332	3.4	-0.7
Ivano-Frankovsk	2.8	349	3.6	0.8
Kirovograd	2.4	194	2.0	-0.4
Lugansk	5.4	467	4.8	-0.6
L'vov	5.3	642	6.6	**1.3**
Nikolaev	2.6	230	2.3	-0.3
Î dessa	5.0	515	5.3	0.3
Poltava	3.4	354	3.6	0.2
Rovno	2.2	247	2.5	0.3
Sumy	2.8	274	2.8	0.0
Ternopol	2.3	250	2.6	0.3
Kharkov	6.1	545	5.6	-0.5
Kherson	2.4	232	2.4	0.0
Khmelnitskiy	2.9	301	3.1	0.2
Cherkassy	2.9	245	2.5	-0.4
Chernovtsy	1.8	200	2.0	0.2
Chernigov	2.7	260	2.7	0.0
TOTAL	100	9792	100.2	

As seen from presented tables, maximal error of households sample makes: for distribution according to the households size - 1.6%, sex - 1%, age - 2.6%, type of populated area - 0.8%, and oblasts - 1.3%. Thus it is possible to suggest that the error in household sample in this survey research does not exceed 2.6%.

3. INTERVIEWERS TRAINING

3.1. The structure of interviewers network

KIIS has the all-Ukrainian network of interviewers consisting of 27 brigades situated in 24 oblasts of Ukraine and the Crimea. Coordination of field work on places is made by team leaders of interviewers groups, general management - by the coordinator of interviewers network in Kiev.

Besides, network of field control, which does not cross with the interviewers network is also within the KIIS structure. Among the tasks of field control network is monitoring the interviewers job by way of recurrent attending a selected number of addresses where the polling was held.

Teams of interviewers are mainly situated in oblast centres of Ukraine, they hold polls not only in their city, but also in all populated areas of the given oblast.

In total, the network of interviewers accounts for more than 300 interviewers - their quantity varied depending on the scale of survey research. Practically all of them participated in this research.

3.2. Interviewers training

The following scheme of training was used in this survey research:

1) Centralised training of representatives of oblast interviewers teams in 4 regional centres (Centre, East, South, West) by KIIS instructors.

2) Local training of each oblast interviewers by team leader and/or another team representative who was trained in a regional centre.

Training in each regional centre was held during 2 days, local training - during 1 day.

In total, the stage of training interviewers started in May, 13 with training the Kiev group of interviewers and finished in May, 29 with training teams from the Southern oblasts of Ukraine (See the Chart of training in SUPPLEMENT).

Training consisted of several stages:

Stage 1. Theoretical training.

1. Summary of the main rules of holding an interview.

2. Discussion of specific character of the given research.

3. Rules of selecting respondents for the 3 types of sampling for the given research.

4. Discussion of terminology for the given research (questionnaire of the given research contains a number of economical terms that are not typical for sociological surveys).

5. Discussion of questions from the questionnaire and rules of its filling.

6. Rules of filling other KIIS documents.

II stage. Independent work.

Filling trial questionnaires and writing final test (including assignments on checking knowledge of general rules of holding the interview and rules of filling a questionnaire for given survey research) by interviewers at the end of training.

III stage. Check of trial questionnaires and tests.

Discussion with instructor of correct answers and typical mistakes of interviewers during the fulfilment of independent assignment.

It is also worth noting that on holding polling according to address sampling, the team leaders of many oblasts enlarged their groups of interviewers, as the size of work at this stage was higher then we primarily supposed. Team leaders independently trained new interviewers, tested them, but they didn't send their tests to the KIIS.

At the large stage of training the instructors checked 200 tests, which makes 70% of the total number of interviewers participating in training. Only 3 interviewers were not admitted for participation in survey research after the check. Maximal number of mistakes made by the interviewers in testing - 10%.

4. FIELD WORK: PROBLEMS OF INTERVIEWERS IN WORK WITH QUESTIONNAIRES.

Field stage in households and “street” business polling was held in March 29 - May 4.

Field stage in polling enterprises according to the lists of State Committee for Statistics was held in April 4 - June 8. In April 29,7% of necessary number of interviews were collected (46.3% of

total number of interviews held), in May - 35.5% of necessary number of interviews (46% of total numbers of interviews held), in June-July - 6% of necessary number of interviews (8% of interviews held).

4.1. Making field stage of research on address sampling:

The problems, which occurred during the polling of enterprises on addresses of State Committee of Ukraine for Statistics, were primarily connected with the fact that a lot of enterprises didn't exist on given addresses. Thus, the search for enterprises and holding the polling needed the attraction of additional number of interviewers and delayed the terms of field works.

Table 4.1.1. gives statistics of the field stage of given type of research:

Table 4.1.1.

		1-10 employees	11-50 employees	51-250 employees	251+ employees	TOTAL
1.	Number of enterprises in the lists (main list, additional etc.)	5593	2802	2791	2814	14000
2.	Number of the checked addresses	4766	1862	1779	1709	10116
2.1.	Enterprise should not be interviewed	354	46	75	77	552
2.2.	Not existed enterprises	2610	560	336	164	3670
2.3.	Found enterprises	1802	1256	1368	1468	5894
2.3.1.	Number of completed interviews	1187	895	934	961	3977
3.	request to interviewers for number of completed interviews	2000	1000	1000	1000	5000
	RESPONSE RATE (share of completed interviews from all found enterprises)	**65.9**	**71.3**	**68.3**	**65.5**	**67.5**
	Share of non-taken interviews in checked lists	75.1	51.9	47.5	43.8	60.7

As seen from the table, the biggest share of non-taken interviews falls on the list of enterprises of the 1st group of employment, the smallest - on the 4th group. The next table presents the reasons of not taking interview:

Table 4.1.2.

		1-10 employees	11-50 employees	51-250 employees	251+ employees	TOTAL
1	Enterprise should not be interview	9.89%	4.76%	8.88%	10.29%	8.99%
2	Not existed enterprise	72.93%	57.91%	39.76%	21.93%	59.78%
3	Refused enterprise	17.18%	37.33%	51.36%	67.78%	31.23%

Thus, during the field stage of research 3977 questionnaires were collected. In final data file there are 3904 interviews, 73 questionnaires were rejected on the following reasons:

- 66 questionnaires on Kharkov were collected above the planned quantity.
- 1 questionnaire is rejected at the stage of coding information because of a large share of Non-Answers.
- 6 questionnaires are moved off the data file on field control results.

4.2. Making field works of households polling

The task for sampling included 202 populated areas - 91 cities, 26 city-type settlements, 85 villages.

The total number of households which participated in survey research was 4008. 929 households were selected for polling but didn't take part in it. The reasons of non-participants of households in polling were the following:

1) Absence from home of one of household members during 3 visits of an interviewer - 481 households (51.7%).

2) Refusal to open the door for an interviewer - 113 households (12.2%).

3) Refusal from participation in polling after the explanation by an interviewer of the purpose of his visit - 178 households (19.2%).

4) Another reason of interview failure - 157 households (16.8%).

Thus, the share of household which answered the polling questions made 81.2%.

In total, in 4008 households live 9792 persons aged 15 and older. According to the filled household registers 5320 persons of this total number are not employed for the moment of polling. Of them 13.3% are pupils and students, 8.9% are engaged in household economy, 55,8% are pensioners, 19.4% are unemployed, and 2.7% are not employed because of another reason.

Of 4472 employed household members 172 persons combine work on hire and self-employment, 3813 have a work on hire, and 831 persons are self-employed (of them 172 persons combine both activities). Of all persons to be polled were not questioned: working on hire - 546 persons, self-employed - 185 persons.

The reasons of non-participation of household members in polling distributed in the following way:

Table 4.2.1.

	Workers on hire		Self-employed	
	N	%	N	%
Another household member working at the same enterprise is questioned	255	46.6	70	37.6
Refusal from participation in an interview (the register is filled from words of another household member)	40	7.4	37	20.1
Long-term absence of the respondent	188	34.4	46	24.8
The respondent is ill	23	4.2	5	2.7
Broken interview	0	0.0	2	1.3
Another reason	40	7.4	25	13.4
TOTAL	546	100	185	99.9

So, the share of household members who gave answers is calculated in the following way:

1) as the first reason of non-participation in polling is that the given respondent didn't have to be questioned at all, the number of interviews held had to make 3558+761=3629.

2) as in 4008 households 3629 interviews had to be made, in 929 non-questioned households there must be held correspondingly 841 interviews.

3) So, the total number of potential respondents made 3629+841=4470 persons. Of them questioned – 3267+646=3913 persons. Thus, the share of those who gave answers is: 3913*100% / 4470 = **87.5%**

Nevertheless it is worth noting that the given indicator is sooner of formal (declarative) nature - in real terms the number of respondents is much lower. Unfortunately, we will not be able to evaluate the real response rate, as we do not know the share of household members under survey, who did not give a sincere answer about their activity. We can only assume that a certain part of respondents concealed that they are engaged in entrepreneurship.

4.3. Holding field works on random sampling

Street polling of "visible" businesses was held in 82 populated areas, of them: in 25 oblast centres, 43 other cities (towns), and in 14 urban-type communities (UTC). In total, during the field stage 991 interviews with subjects of entrepreneurship were collected. As the number of "visible" enterprises on the territory postal districts research is different, and in some territories there are no such enterprises at all, the number of collected interviews is a little bit lower, then it was planned - 991 instead of 1000.

Besides the interviews, the information on 3026 randomly selected "visible" enterprises is collected in this research.

Òàble 4.3.1. Distribution of "visible" enterprises according to the number of employed in representative sampling (N=3026) and among polled enterprises (N=991)

Number of employed	All "visible" enterprises		All interviewed enterprises	
	Frequency	%	Frequency	%
1 employed	1256	41.5	583	58.8
2-10 employed	1431	47.3	338	34.1
11-5- employed	259	8.6	58	5.9
51-250 employed	41	1.4	10	1.0
251 and more employed	18	.6	2	.2
Refused to answer	20	.6	0	0
TOTAL	3026	100.0	991	100.0

Among the enterprises selected by the interviewers for polling 491 enterprises refused from participating in the interview.

Besides 991 collected questionnaires, 5 questionnaires were added to the data file from household sampling. These questionnaires were obtained under unwrapping of polling chain in the places of households dwelling. 5 addresses selected for polling were the addresses of offices, and not households.

Problems in holding such type of survey:

Among the respondents who answered the questionnaire questions there were many people engaged in individual commercial activity. Among these people there were many pensioners, who had to trade in the street in order to make both ends meet. It was rather difficult for these people to answer some questions from questionnaire, as these questions were sooner designed for enterprises and were difficult for understanding to ordinary old women who do not consider themselves entrepreneurs. Answer "Hard to say" is often met in such questionnaires.

5. FIELD CONTROL

Monitoring of KIIS polling network interviewers work is executed in several directions:

- quality control for making sampling (unwrapping polling network, developmental work on addresses in case of address sampling, etc.)
- monitoring the fact of holding an interview
- quality control for holding an interview.

Tool for quality control of polling chain unwrapping in KIIS standard sampling is the interviewer's diary. At this stage the team leaders checked on spot (in the provinces) the correctness of filling the diary, and also the correctness of building chain for selection of

households and the route of selecting “visible” enterprises for street polling. After passing the questionnaires and interviewers diaries to Kiev the majority of diaries was checked, and if there were any doubts the questionnaires of the given questioning chain were put on control. During the field monitoring of address sampling both - the enterprises where the questioning was held, and the addresses where the questioning was not held (quality control of addresses check - whether the reasons of not taking interview were valid) were controlled.

In this research field control was held in all oblasts of Ukraine. Practically all interviewers who took part in polling were controlled. The share of interviews which ha to be controlled makes 10% of the total number.

The tool for controlling the fact and quality of polling was the specially developed inspector questionnaire.

In household polling 441 households participating in polling from 15 oblasts of Ukraine were controlled. In 6 households the violations of selecting rules were observed, and all questionnaires of these households were removed from data file.

Numbers of questionnaires of these households:

À) Registers of households employment - ¹ ¹ 2424, 2725, 2727, 2733, 2826, 3515.

B)

In street polling 10 polling chains in different oblasts of Ukraine were controlled. In these chains 297 enterprises introduced into the route map of businesses (representative list of enterprises) and 100 enterprises that were polled by questionnaire were recorded (fixed). All information on 10 polling chains was confirmed completely, no violations were observed.

In address sampling of enterprises 350 enterprises participating in questioning and situated in all oblasts of Ukraine were controlled on subject of fact and quality of polling holding. The fact of polling was not confirmed at 6 enterprises (N 4209, 4987, 4988, 4993, 5015, 5020), the questionnaires of which were removed from data file. 153 enterprises of those, which could not be found by interviewers on a given address, were checked. This control showed that the results of address check were not falsified - there are no such enterprises on given addresses.

6. PREPARING DATA FILE ON ELECTRONIC COMPUTER

Coding and editing of questionnaires was held in following directions:

- registration of obtained questionnaires, giving them unique codes (numeration), coding sampling characteristics - oblasts (spheres) of polling, populated area (codes of settlements are given in SUPPLEMENT), check of correctness of filling chapters for interviewers, in other words - features, information of which can be compared with or reconstructed by interviewers' diaries;

- check of correctness of questionnaires' filling: observing transitions correct understanding of the polling logic, analysis of interviewers' mistakes.

8 coders participated in coding and editing questionnaires; a team of operators of 10 persons did input of questionnaires.

7. LOGICAL CONTROL

Logical control of the data quality is a computer program, which checks co-ordination of respondents' answers to certain questions of the questionnaire. For example, if the enterprise has the state form of ownership, its director can't be the owner of this enterprise (coordination of questions A5 and A7), etc.

In total, more than 200 logical conditions were formulated in the logical control program, and about 30% of questionnaires were checked.

At each stage of control on formulated logical conditions the questionnaires, which seemed suspicious about possible mistakes in them were selected. Hard copy of each questionnaire was checked and in case of necessity corrections were introduced into the questionnaire and data file.

Besides, the control of outlays of quantitative variables (e.g., deviating values of working day duration, etc.) was held.

8. WHAT INCLUDES THE ELECTRONIC ARCHIVES (FILES) OF RESEARCH DATA:

A. Household Survey

1. ***Data file "ROSTER.SAV" – data from the roster of the Household's Employment***

2. ***Data file "HIRED.SAV" - data about enterprises , where household's members are hired employees***

3. ***Data file "SELF.SAV" - data about activity of self-employed household's members***

B. Street Survey

4. ***Data file "RANDOM.SAV" – data about smallest street enterprises***

5. ***Data file "ROUT.SAV" - data about employment of all street enterprises***

C. Register Survey

6. ***Data file "ADDRESS.SAV" – data about registered enterprises from GosComStat's lists***

And Combined data file "COMBINED.SAV"